한국통계학회논문집 제3권 1호
1996년 4월 pp. 145-153

A Modified Chen-Wolfe Procedure for Comparing Umbrella Pattern Treatment Effects with a Control in a One-way Layout

Dong Hoon Lim[1], Soo Taek Kim[2], Joong Yang Park[3]

Abstract

Nonparametric tests for comparing umbrella pattern treatment effects with a control in a one-way layout were studied in Chen and Wolfe (1993). In this paper we propose a modification that improves the power of the Chen-Wolfe test. The results of a Monte Carlo power study are discussed.

1. Introduction

Nonparametric procedures for comparing several treatments with a control in a one-way layout have been studied. For example, Dunn (1964) suggested a multiple rank test for the general setting in which no information about the pattern of treatment effects is available. Shirley (1977) considered a nonparametric version of Williams' (1971,1972) test for the situation where the experimenter has the information that if there were a response to the substance the treatment effects would be monotonically ordered. Williams (1986) further gave a modification of Shirley's procedure. However, monotonicity of dose-response relationships is far from universal. For example, many therapies for humans become counter productive at high doses. In this case, an increasing dose-response relationship with a downturn in response at high doses is anticipated. Since this corresponds to an up-down ordering of the treatment effects, they are said to follow an umbrella pattern [see, for example, Mack and Wolfe (1981)]. Chen and Wolfe (1993) have recently considered comparing umbrella pattern treatment effects with a control. In the Chen-Wolfe test their ranks of observations are unchanged as each treatment level is tested. In this paper we develop a more powerful test by reranking the observations at each stage of the sequential test procedure, after excluding all observations at those treatment levels at which significant evidence of an effect has already been established.

1) Assistant Professor, Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.
2) Professor, Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.
3) Professor, Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.

Suppose that $X_{i1}, \ldots, X_{in_i}$, $i=0,1,\ldots,k$, are $k+1$ independent random samples from populations with continuous distribution functions $F_i(x)=F(x-\theta_i)$, $i=0,1,\ldots,k$, respectively. The zero population ($i=0$) is the control and the other k populations are treatments. Under the prior belief of $\theta_1 \le \ldots \le \theta_p \ge \ldots \ge \theta_k$ for some p, we consider the problem of deciding $\theta_i > \theta_0$ when the peak of the umbrella p is known or unknown.

In the following Section 2 the Chen-Wolfe procedure is briefly discussed. In Section 3 the modification of the Chen-Wolfe procedure is proposed, which increases its power. In Section 4 a detailed example illustrating the modified procedure is also given. In Section 5 we present the results of a Monte Carlo simulation investigation of the relative powers of the Chen-Wolfe test and its modification.

2. Chen-Wolfe Procedure

Let R_{ij} be the rank of X_{ij} among the $N=\sum_{i=0}^{k} n_i$ observations and $\overline{R}_i=\sum_{j=1}^{n_i} R_{ij}/n_i$ be the average rank of the ith sample. If ties occur in the rankings, average ranks are used. Suppose that the peak of the umbrella is known to be at group $p(1 \le p \le k)$. Further, let $\widehat{R}_1^{(p)} \le \ldots \le \widehat{R}_p^{(p)} \ge \ldots \ge \widehat{R}_k^{(p)}$ be the isotonic regression of $\overline{R}_1, \ldots, \overline{R}_k$ under the restriction $\theta_1 \le \ldots \le \theta_p \ge \ldots \ge \theta_k$. For deciding $\theta_i > \theta_0$, Chen and Wolfe (1993) suggested a multiple test based on the statistics

$$T_{p:i}=(\widehat{R}_i^{(p)}-\overline{R}_0)[V(1/n_i+1/n_0)]^{-1/2}, \quad i=1,\ldots,k, \tag{2.1}$$

with $V=N(N+1)/12-T$, where T denotes the correction for ties in which any group of t tied ranks contributes $(t^3-t)/(12(N-1))$ to T. Note that the statistics $T_{k:i}$ are actually Shirley'(1977) statistics for comparing ordered treatment effects with a control.

Let $t(\alpha;k,p)$ be the upper αth percentile of the distribution of $T_{p:p}$ under the null hypothesis that all the θ_i's are equal. If $T_{p:p} \ge t(\alpha;k,p)$, the Chen-Wolfe procedure then claims that at least the dose p is better than the control. They then suggested starting from both the doses $p-1$ and $p+1$ to search for those doses which were more effective than the control. If $T_{p:p-1} < t(\alpha;k,p)$ and $T_{p:p+1} < t(\alpha;k,p)$, they concluded that there was evidence for a response only at dose p. If, however, $T_{p:p-1} \ge t(\alpha;k,p)$(and/or $T_{p:p+1} \ge t(\alpha;k,p)$), they claimed that there was evidence for a response at doses p and $p-1$(and/or $p+1$) and then suggested testing for a response at dose $p-2$ (and/or $p+2$). This procedure is continued until dose levels u and v are obtained for which $T_{pu-1} < t(\alpha;k,p)$ and $T_{pv+1} < t(\alpha;k,p)$, where $1 \le u < p < v \le k$. Finally, they concluded that there was evidence for a response at doses $u, \ldots, v$.

For the more general setting in which no information concerning the location of the peak group is available, Chen and Wolfe employed the method suggested by Mack and Wolfe (1981) to obtain an estimate of the unknown peak. Let U_{ij} be the usual Mann-Whitney statistic corresponding to the number of observations in sample j that exceed observations in sample i and let $Z_t = \sum_{i \neq t}^{k} U_{it}$, $t = 1, \ldots, k$. Then if we assume that $n_1 = \cdots = n_k$ and let $r(1 \leq r \leq k)$ equal the number of populations tied for having the largest Z_j sample value, we set

$$\chi_t = \begin{cases} 1/r, & \text{if the } t\text{th population is among those tied for the largest } Z_j \text{ value} \\ 0, & \text{otherwise.} \end{cases}$$

Usually r will equal one (i.e., there will be only one population, say the tth one, which attains the maximum Z_j sample value). However, since there is a positive probability that two or more of the variables $Z_1, \cdots, Z_k$ will be tied for the maximum, the above definition of χ_t is necessary.

They proposed a multiple test procedure similar to that described above, but employing the test statistics

$$T_{\hat{p}:i} = \sum_{t=1}^{k} \chi_t T_{t:i}, \quad i = 1, \cdots, k, \tag{2.2}$$

and critical values $t(\alpha, k, \hat{p})$ of the test based on $T_{\hat{p}:i}$, to determine which treatments are better than the control.

Unfortunately, critical values for $T_{\hat{p}:i}$ is not presented in the Chen-Wolfe's paper. In Table 2 we present simulated critical values obtained by estimating the cumulative distribution function of $T_{\hat{p}:i}$ using uniform random numbers.(For details concerning the estimation of the distribution function, see Chen and Wolfe (1993), Section 4).

3. A Modified Chen-Wolfe Procedure

We begin by defining statistics for comparing k' ($< k$) treatment groups $u, \ldots, v$(excluding the significant treatment levels) with a control. Let $N_{k'} = n_0 + \sum_{i=u}^{v} n_i$denote the total number of observations in these groups and $V_{k'} = N_{k'}(N_{k'}+1)/12 - T_{k'}$, where $T_{k'}$ denotes the correction for ties in which any group of t tied ranks contributes $(t^3 - t)/(12(N_{k'} - 1))$ to $T_{k'}$. Suppose that the peak for these groups is known to be at group $m(u \leq m \leq v)$. For deciding $\theta_m > \theta_0$ under the restriction $\theta_u \leq \ldots \leq \theta_m \geq \ldots \geq \theta_v$ the modification to (2.1) is

$$\hat{T}_{k':m} = (\widehat{R_m}^{(m)} - \overline{R}_0)[\ V_{k'}(1/n_m + 1/n_0)]^{-1/2}. \tag{3.1}$$

In particular, the test based on $\hat{T}_{k:k}$ is Williams (1986) test for comparing ordered treatment effects with a control. For deciding $\theta_p > \theta_0$ at all $k+1$ treatments with the peak p, the test statistic $T_{p:p}$ in (2.1) with $i=p$ is identical with $\hat{T}_{k:p}$ in (3.1) with $k'=k$ and $m=p$. Usually, the (3.1) will be greater than (2.1) because $V_{k'}$ is substantially less than V.

The procedure is as follows. At first, we apply the statistic $\hat{T}_{k:p}$ on all $k+1$ groups $0, 1, \ldots, k$ in the case that the peak is known to be at group p. If $\hat{T}_{k:p} \geq t(\alpha, k, p)$, we conclude that group p is better than the control. We then rerank all observations at the remaining k groups $0, 1, \ldots, p-1, p+1, \ldots, k$ with the peak, say, m, after excluding the significant group p. The $\hat{T}_{k-1:m}$ is calculated in the same way and compared to its critical value $t(\alpha, k-1, m)$. This process continues until a nonsignificant treatment effect is found.

The test for our problem in the case that the peak m is unknown is based on

$$\hat{T}_{k':\hat{m}} = \sum_{t=u}^{v} \chi_t T_{k':t}, \tag{3.2}$$

where χ_t's are the random variables used in defining $T_{\hat{p}:i}$ (2.2). The procedure based on $\hat{T}_{k':\hat{m}}$ can be used in the same manner as the procedure based on (3.1), but estimating the unknown peak at each stage of the sequential test procedure and employing the critical value $t(\alpha, k', \hat{m})$.

4. An Example

Consider the data in Table 1 analyzed in Section 5 of the Chen-Wolfe's paper, in which there are 5 dose levels and a control.

Table 1
Revertant colonies for Acied Red 114, TA98, hamster liver activation

Dose (ug/ml)					
0	100	333	1,000	3,333	10,000
23	27	28	41	28	16
22	23	37	37	21	19
14	21	35	43	30	13

For detecting dose levels that are better than the zero-dose control, the modified Chen-Wolfe statistic $\hat{T}_{k:m}$ (3.2) is applied to the data of Table 1. We first apply the statistic to all six dose levels including the control. For these data, the peak $\hat{p}$=3 (1,000 ug/ml) is obtained (i.e., $\chi_3=1$, $\chi_1=\chi_2=\chi_4=\chi_5=0$ and $\overline{R}_0=5.83$, $\overline{R}_1=8.0$, $\overline{R}_2=13.67$, $\overline{R}_3=16.83$, $\overline{R}_4=10$ and $\overline{R}_5=2.67$ are obtained. To test for an effect at the 1,000 ug/ml dose, the correction for ties is

$$\{4(2^3-2)\}/\{(12)(17)\}=.1176$$

and

$$\hat{T}_{5:3}=(16.83-5.83)[\ \{(18)(19)/12-.1176\}(2/3)]^{-1/2}=2.529.$$

For α=0.05, the approximate critical value is $t(\alpha;5,3)$ =2.141 using Table 2, so there is a significant effect at the 1,000 ug/ml dose.

Now, we rerank the data for the remaining five dose levels excluding all the observations at 1,000 ug/ml dose. This gives mean ranks of 5.83, 8.0, 13.5, 10 and 2.67 for the 0, 100, 333, 3,333 and 10,000 ug/ml doses, respectively. For these data, we find the estimated peak group to be $\hat{p}$=2 (333 ug/ml). To test for an effect at the 333 ug/ml dose, the correction for ties is

$$\{3(2^3-2)\}/\{(12)(14)\}=.1071$$

and

$$\hat{T}_{4:2}=(13.5-5.83)[\ \{(15)(16)/12-.1071\}(2/3)]^{-1/2}=2.1062.$$

For α=0.05, the approximate critical value is $t(\alpha;4,2)$ =2.100, so there is a significant effect at the 333 ug/ml dose.

Note that the Chen-Wolfe's method here gives statistic

$$T_{3:2}=(13.67-5.83)[\ \{(18)(19)/12-.1176\}(2/3)]^{-1/2}=1.802,$$

which, being less than 2.141, is not significant at the 5% level.

Reranking the remaining 0, 100, 3,333 and 10,000 ug/ml doses gives mean ranks of 5.83, 8.0, 9.5 and 2.67, respectively, and we obtain the peak $\hat{p}$=2 (3,333 ug/ml). To test for an effect at 3,333 ug/ml dose, the correction for ties is

$$\{2(2^3-2)\}/\{(12)(11)\}=.0909$$

and

$$\hat{T}_{3:2}=(9.5-5.83)[\ \{(12)(13)/12-.0909\}(2/3)]^{-1/2}=1.2510,$$

which is below the 5% critical value of 2.038. So, we conclude, at the 5% significance level, that the dosages between 333 ug/ml and 1,000 ug/ml are more effective than the zero-dose control.

5. Monte Carlo Power Study

To investigate the relative powers of the Chen-Wolfe test $T(p)$ and $T(\hat{p})$, and the modified Chen-Wolfe test $\tilde{T}(p)$ and $\tilde{T}(\hat{p})$ for both cases where the peak of the umbrella is known or unknown, we conducted a Monte Carlo power study. The study was performed for k=4 populations with $n_0 = n_1 = \ldots = n_k = 5$ observations per sample and for a variety of different umbrella pattern treatment effects. The designed configurations of the treatment effects correspond to values of $\theta_{10} = \theta_1 - \theta_0, \ldots, \theta_{k0} = \theta_k - \theta_0$. For each of these settings, the IMSL routine GGNML and GGEXN were used to generate normal and exponential random numbers. In each case, we used 10,000 replications to obtain the various power estimates. The approximate critical values for peak-known tests were used from Chen and Wolfe (1993), and for peak-unknown tests were used from Table 2. The power π_{i0} represents the probability of declaring $\theta_i > \theta_0$, $i = 1, \ldots, k$. The simulated power estimates for these tests considered in this study are given in Table 3. As expected for the peak group-control comparison on the total data set, the test $T(p)$(or $T(\hat{p})$) has the same power as the test $\tilde{T}(p)$(or $\tilde{T}(\hat{p})$). However, for the detection of effects at the other treatments excluding the peak proup, the test $\tilde{T}(p)$(or $\tilde{T}(\hat{p})$) is better than the test $T(p)$ (or $T(\hat{p})$). This is because the variance which is used the statistic $\tilde{T}(p)$(or $\tilde{T}(\hat{p})$) is substantially less than that in the statistic $T(p)$(or $T(\hat{p})$). As is illustrated by the example in Section 4, however, the test $\tilde{T}(\hat{p})$ has the cumbersome process in estimating the unknown peak at each stage of the procedure.

References

[1] Chen,Y.I.and Wolfe,D.A.(1993). Nonparametric procedures for comparing umbrella pattern treatment effects with a control in a one-way layout.*Biometrics,* 49, 455-465.

[2] Dunn,O.J.(1964). Multiple comparisons using rank sums. *Technometrics,* 6, 241-252.

[3] Mack,G.A. and Wolfe,D.A.(1981). K-sample rank tests for umbrella alternatives. *Journal of the American Statistical Association,* 76, 175-181.

[4] Shirley,E.(1977). A nonparametric equivalent of Williams' test for contrasting increasing dose levels of a treatment. *Biometrics,* 33, 386-389.

[5] Williams,D.A.(1971). A test for differences between treatment means when several dose levels are compared with a zero-dose control. *Biometrics,* 27, 103-117.

[6] Williams,D.A.(1972). The comparison of several dose levels with a zero-dose control.*Biometrics,* 28, 519-531.

[7] Williams,D.A.(1986). A note on Shirley's nonparametric test for comparing several dose levels with a zero-dose control. *Biometrics,* 42, 183-186.

Table 2 Approximate critical values for $T_{\hat{p}:i}$

		k = No. of treatments(excluding a control)			
n	Level	2	3	4	5
	0.01	2.236	2.491	2.556	2.600
3	0.05	1.938	2.038	2.100	2.141
	0.10	1.640	1.812	1.826	1.912
	0.01	2.451	2.525	2.630	2.650
4	0.05	1.937	2.005	2.151	2.200
	0.10	1.569	1.764	1.853	1.900
	0.01	2.475	2.566	2.664	2.730
5	0.05	1.909	2.031	2.148	2.227
	0.10	1.626	1.764	1.848	1.904
	0.01	2.433	2.613	2.689	2.767
6	0.05	1.893	2.041	2.164	2.219
	0.10	1.568	1.755	1.869	1.918
	0.01	2.498	2.599	2.686	2.745
7	0.05	1.895	2.047	2.139	2.222
	0.10	1.594	1.754	1.826	1.939
	0.01	2.510	2.585	2.695	2.768
8	0.05	1.909	2.047	2.138	2.232
	0.10	1.591	1.732	1.839	1.911
	0.01	2.465	2.573	2.692	2.787
9	0.05	1.901	2.058	2.136	2.232
	0.10	1.574	1.745	1.830	1.933
	0.01	2.515	2.601	2.715	2.791
10	0.05	1.905	2.027	2.148	2.202
	0.10	1.575	1.741	1.856	1.895

Table 3 Power estimates for $k=4$, $n_0=n_1=\ldots=n_4=5$ and $\alpha=0.05$

(a) Normal

θ_{10}	θ_{20}	θ_{30}	θ_{40}		$T(p)$	$\tilde{T}(p)$	$T(\hat{p})$	$\tilde{T}(\hat{p})$
0	0	.5	1	π_{10}	--	--	--	--
				π_{20}	--	--	--	--
				π_{30}	.112	.148	.047	.097
				π_{40}	.376	.376	.227	.227
.5	1	1	1.5	π_{10}	.060	.092	.014	.092
				π_{20}	.241	.383	.099	.287
				π_{30}	.383	.442	.199	.326
				π_{40}	.705	.705	.522	.522
0	.5	1	0	π_{10}	--	--	--	--
				π_{20}	.099	.150	.047	.097
				π_{30}	.364	.364	.233	.233
				π_{40}	--	--	--	--
.5	1	1.5	.5	π_{10}	.067	.163	.026	.163
				π_{20}	.303	.398	.175	.295
				π_{30}	.675	.675	.522	.522
				π_{40}	.089	.207	.041	.139
0	1	.5	0	π_{10}	--	--	--	--
				π_{20}	.360	.360	.230	.230
				π_{30}	.100	.152	.046	.100
				π_{40}	--	--	--	--
1	1.5	.5	.5	π_{10}	.294	.397	.168	.296
				π_{20}	.673	.673	.516	.516
				π_{30}	.113	.213	.050	.144
				π_{40}	.039	.169	.012	.169
1	.5	0	0	π_{10}	.380	.380	.226	.226
				π_{20}	.113	.145	.045	.093
				π_{30}	--	--	--	--
				π_{40}	--	--	--	--
1.5	1	1	.5	π_{10}	.698	.698	.520	.520
				π_{20}	.378	.431	.197	.320
				π_{30}	.237	.386	.095	.288
				π_{40}	.061	.090	.013	.090

(b) Exponential

θ_{10}	θ_{20}	θ_{30}	θ_{40}		$T(p)$	$\tilde{T}(p)$	$T(\hat{p})$	$\tilde{T}(\hat{p})$
0	0	.5	1	π_{10}	--	--	--	--
				π_{20}	--	--	--	--
				π_{30}	.182	.215	.083	.142
				π_{40}	.547	.547	.370	.370
.5	1	1	1.5	π_{10}	.114	.202	.041	.202
				π_{20}	.409	.575	.236	.475
				π_{30}	.545	.611	.371	.505
				π_{40}	.837	.837	.722	.722
0	.5	1	0	π_{10}	--	--	--	--
				π_{20}	.177	.224	.089	.149
				π_{30}	.525	.525	.364	.364
				π_{40}	--	--	--	--
.5	1	1.5	.5	π_{10}	.136	.262	.063	.262
				π_{20}	.491	.600	.346	.501
				π_{30}	.822	.822	.714	.714
				π_{40}	.148	.339	.076	.247
0	1	.5	0	π_{10}	--	--	--	--
				π_{20}	.528	.528	.374	.374
				π_{30}	.169	.221	.085	.145
				π_{40}	--	--	--	--
1	1.5	.5	.5	π_{10}	.485	.605	.337	.504
				π_{20}	.831	.831	.723	.723
				π_{30}	.182	.340	.093	.247
				π_{40}	.080	.268	.026	.268
1	.5	0	0	π_{10}	.543	.543	.365	.365
				π_{20}	.190	.219	.090	.146
				π_{30}	--	--	--	--
				π_{40}	--	--	--	--
1.5	1	1	.5	π_{10}	.837	.837	.716	.716
				π_{20}	.546	.618	.373	.505
				π_{30}	.406	.573	.237	.479
				π_{40}	.110	.200	.039	.200